



Universidade Técnica de Lisboa  
Instituto Superior de Economia e Gestão



Mestrado em: Economia Monetária e Financeira

## **Aplicação Empírica da *Realized Volatility* ao Índice PSI20**

Júlio António Rocha Delgado

Orientação: Doutor João Carlos Henrique da Costa Nicolau

Júri:

Presidente: Doutor Daniel de Assunção Müller, professor catedrático do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa;

Vogais: Doutor José Joaquim Dias Curto, professor auxiliar da Escola de Gestão do Instituto Superior de Ciências do Trabalho e da Empresa;

Doutor João Carlos Henrique da Costa Nicolau, professor auxiliar do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa.

Junho de 2005



Universidade Técnica de Lisboa  
Instituto Superior de Economia e Gestão



Mestrado em: Economia Monetária e Financeira

## **Aplicação Empírica da *Realized Volatility* ao Índice PSI20**

Júlio António Rocha Delgado

Orientação: Doutor João Carlos Henrique da Costa Nicolau

Júri:

Presidente: Doutor Daniel de Assunção Müller, professor catedrático do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa;

Vogais: Doutor José Joaquim Dias Curto, professor auxiliar da Escola de Gestão do Instituto Superior de Ciências do Trabalho e da Empresa;

Doutor João Carlos Henrique da Costa Nicolau, professor auxiliar do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa.

Junho de 2005

## Algumas Notações e Abreviações

$\sigma_t^2$	variância (i.é. <i>realized volatility</i> )
$\ln \sigma_t^2$	logaritmo da variância
$r_t$	retorno diário
$r_t/\sigma_t$	retorno diário estandardizado
$p_t$	logaritmo do preço
$m$	número de observações intra-diários
$\Delta$	frequência amostral
ARCH	<i>autoregressive conditional heteroscedasticity</i>
ARFIMA	<i>autoregressive fractionally integrated moving average</i>
JB	Jaque-Bera
DFA	Dickey-Fuller Aumentado
FAC	função de autocorrelação parcial
LB	Ljung e Box
RV	<i>realized volatility</i>
VE	volatilidade estocástica
VI	volatilidade integrada
VQ	variação quadrática

## Aplicação Empírica da *Realized Volatility* ao Índice PSI20

Júlio António Rocha Delgado

Mestrado em: Economia Monetária e Financeira

Orientador: Professor Doutor João Carlos Henrique da Costa Nicolau

Provas concluídas em:

### Resumo

Nesta dissertação é feito um estudo sobre o novo método ou procedimento (não paramétrico) de estimação da volatilidade recentemente proposto na literatura, a *Realized Volatility* (RV), obtido pela soma dos produtos cruzados dos retornos de alta-frequência intra-diários. O objectivo principal do estudo consiste em fazer uma aplicação empírica da RV ao índice PSI20, focando sobretudo nos estudos das propriedades das distribuições condicionais e não condicionais, confrontando com os resultados já obtidos na literatura. Considerando duas séries de frequências de 5 e 30 minutos verificamos que ambas as distribuições empíricas da RV são não normais e fortemente enviesado à direita, enquanto que as distribuições marginais do logaritmo da RV são aproximadamente normais, assim como as distribuições dos retornos estandardizados. O logaritmo da RV apresenta uma forte dependência temporal e parece ser bem descrita por um processo de memória longa. Esses resultados são consistentes com os estudos já efectuados. Verificamos ainda que logaritmo da RV exhibe o efeito assimétrico da volatilidade. Considerando a volatilidade como uma variável observável, em vez de latente como nos modelos ARCH (*AutoRegressive Conditional Heteroskedastic*) e de Volatilidade Estocástica (VE), propomos modelar as características dinâmicas do logaritmo da RV através do modelo ARFIMA (*Autoregressive Fractionally Integrated Moving Average*).



**Palavras Chave:** *Realized Volatility* (RV); Variação Quadrática (VQ); Volatilidade Integrada (VI); Dados de alta-frequência; Memória Longa; Modelos ARFIMA

## **Abstract**

In this dissertation, it is done a study about a new (non-parametric) method or procedure to estimate the volatility recently proposed in literatures, *Realized Volatility* (RV), which can be obtained by summing of cross-products of high frequency intraday returns. The main objective of this study is to perform an empirical application of RV to the PSI20 index, focused on the conditional and unconditional properties, and to compare with the results that have appeared in literature. Considering two series of frequencies 5 and 30 minutes we find for both that the distributions of RV are not normal and are highly skewed to the right, while the log of RV are approximately normal, as are the distributions of the standardized returns. The log of RV show strong temporal dependence and appear to be well described by long memory processes. Our results are consistent with the findings in literature. We also find that the log of RV exhibit the asymmetric volatility effect. Considering the volatility as an observable variable, instead of latent as in ARCH (*AutoRegressive Conditional Heteroskedastic*) and Stochastic Volatility (VE) models, we modelled the dynamic characteristics of the log of RV series using ARFIMA (*Autoregressive Fractionally Integrated Moving Average*) models.

**Key Words:** *Realized Volatility* (RV); Quadratic Variation (VQ); Integrated Volatility (VI); High Frequency Data; Long Memory; ARFIMA models



## Índice

<b>1. Introdução .....</b>	<b>8</b>
<b>2. Breve Enquadramento Teórico da <i>Realized Volatility</i> (RV) .....</b>	<b>14</b>
2.1. <i>Variação Quadrática (VQ) e RV</i> .....	14
2.2. <i>Volatilidade Integrada (VI) e RV</i> .....	17
<b>3. Aplicação Empírica ao índice PSI20 .....</b>	<b>19</b>
3.1. <i>Dados e cálculo da RV</i> .....	19
3.2. <i>Propriedades das Distribuições não Condicionais do Retorno e da RV</i> .....	22
3.2.1. Retorno e Retorno Estandarizado .....	22
3.2.2. RV.....	24
3.3. <i>Propriedades das Distribuições Condicionais da RV</i> .....	27
3.3.1. Dependência Temporal e Memória Longa .....	27
3.3.2. Efeito Assimétrico .....	29
3.4. <i>Modelação da RV</i> .....	31
<b>4. Conclusões .....</b>	<b>34</b>
<b>Bibliografia .....</b>	<b>36</b>



## Lista de Tabelas e Figuras

### Tabelas:

Tabela 1: Distribuição não Condicional do Retorno e Retorno Estandarizado.....	22
Tabela 2: Distribuição não Condicional de $\sigma_t^2$ e de $\lambda\sigma_t^2$ .....	25
Tabela 3: Dependência dinâmica de $\lambda\sigma_t^2$ .....	28
Tabela 4: Assimetria de $\lambda\sigma_t^2$ .....	31
Tabela 5: Estimativas do Modelo ARFIMAX .....	33

### Figuras:

Figura 1: Índice PSI20 (1/Janeiro/1998 a 22/Fevereiro/2004).....	20
Figura 2: Retorno Diário do Índice PSI20 ( $r_t$ ) .....	20
Figura 3: <i>Realized Volatility</i> ( $\sigma_t^2$ ) .....	21
Figura 4: Estimativa da Densidade do retorno ( $r_t$ ) <i>versus</i> Densidade Gaussiana.....	23
Figura 5: Estimativa da Densidade de $r_t/\sigma_t$ <i>versus</i> Densidade Gaussiana .....	24
Figura 6: Estimativa da Densidade de $\sigma_t^2$ <i>versus</i> Densidade Gaussiana .....	25
Figura 7: Estimativa da Densidade de $\lambda\sigma_t^2$ <i>versus</i> Densidade Gaussiana .....	26
Figura 8: Logaritmo da <i>Realized Volatility</i> ( $\lambda\sigma_t^2$ ) .....	28
Figura 9: FAC de $\lambda\sigma_t^2$ .....	28
Figura 10: Assimetria de $\lambda\sigma_t^2$ .....	30
Figura 11: FAC Residual .....	33

## **Agradecimentos**

Quero agradecer a todos que, de uma forma directa ou indirecta, contribuíram para a realização deste trabalho.

Ao meu orientador Professor Doutor João Carlos Henrique da Costa Nicolau, pela sua imprescindível contribuição no acompanhamento, na leitura crítica do trabalho, nas sugestões e, sobretudo, nos conhecimentos transmitidos.

Aos meus queridos pais, familiares e de uma forma muito especial à Rosa Brito Delgado pela compreensão e motivação ao longo da elaboração deste trabalho.

Evidentemente todos os erros e omissões devem-se exclusivamente ao autor.



## 1. Introdução

A importância do risco e da incerteza nos mercados financeiros colocam a volatilidade como um elemento privilegiado e de extrema importância de investigação na área de finanças, com especial relevância entre os académicos/investigadores e os agentes económicos em geral que procurem tomar decisões financeiras (especialmente investidores).<sup>1</sup> Uma medida precisa e boa previsão da volatilidade são essenciais na gestão do risco, afectação e selecção de activos, valorização e *hedging* das opções e derivados e muitas outras actividades financeiras. Por conseguinte, nas últimas décadas os estudos sobre modelação, estimação e a previsão da volatilidade, e paralelamente a avaliação da sua performance têm sido um dos objectos de maior investigação teórica e empírica na literatura financeira e econométrica avançada.

Tradicionalmente, até o início da década de 80 a teoria financeira e econométrica clássica centrava na modelação da média condicionada do processo em estudo, considerando a volatilidade constante ao longo do tempo [nomeadamente, as especificações lineares (ARMA ou ARMAX); ver, por exemplo, Murteira *et al.* (1993)], apesar de ser uma hipótese que raramente se verifica na prática. Constatou-se assim que a via de análise mais correcta para analisar o comportamento estatístico de uma sucessão cronológica financeira consiste em modelar não só a média como também a

---

<sup>1</sup> O termo volatilidade, normalmente utilizado nos mercados financeiros para representar diversas medidas de variabilidade associada às variações dos preços, como é o caso da variância (ou desvio padrão) e da covariância (ou correlação), traduz-se muitas vezes como uma medida do risco total do activo financeiro. Mercados de alta volatilidade significam mercados com grandes variações dos preços (e vice-versa). A intensidade do movimento do preço está intimamente ligada ao fluxo e refluxo da actividade dos mercados financeiros e a chegada de informação a nível intra diário. Ao longo do trabalho utiliza-se o termo genérico de volatilidade para representar as diversas medidas acima referenciadas. No entanto, quando for importante apresenta-se o significado preciso de cada um.

variância condicionada. Neste contexto, destaca-se o modelo ARCH (*AutoRegressive Conditional Heteroskedastic*) proposto por Engle (1982), o qual explicita a variância condicional como função das realizações passadas do processo, ou seja a volatilidade é definida como função de variáveis observáveis.<sup>2</sup> Posteriormente, desencadeou o desenvolvimento de um grande número de modelos que tem por base o modelo ARCH, designadamente o modelo GARCH proposto por Bollerslev (1986), EGARCH sugerido por Nelson (1991), ARCH-M de Engle, Lilien e Robinson (1987), entre outros.<sup>3</sup> Mais recentemente tem havido algum interesse nos modelos de VE, introduzido por Clark (1973), onde variância é função não só dos valores passados do processo como também de flutuações estocásticas, sendo a volatilidade uma variável latente [ver por exemplo Taylor (1986)]. Da mesma forma, têm merecido muita atenção na literatura a análise da volatilidade implícita dos preços das opções e derivados, destacado por Hull e White (1987).

Como alternativa a estes modelos foi recentemente proposto na literatura um novo método ou procedimento (não paramétrico) de estimação da volatilidade – a *Realized Volatility* (RV) [ver Andersen e Bollerslev (1998), Andersen, Bollerslev, Diebold e Labys (doravante ABDL) (2003) e Barndorff-Nielsen e Shephard (2001, 2002)].<sup>4</sup> Basicamente, o método consiste na acumulação dos produtos cruzados dos retornos

---

<sup>2</sup> A Real Academia de Ciências da Suécia atribuiu recentemente o Prémio Nobel da Economia 2003 ao americano Robert F. Engle juntamente com o britânico Clive W.J. Granger pelos seus contributos na área dos métodos estatísticos de análise de séries temporais económicas e financeiras.

<sup>3</sup> Para maiores desenvolvimentos, ver por exemplo, Bera e Higgins (1993), Bollerslev, Engle e Nelson (1994), Shephard (1996), Gouriéroux (1997), Hamilton (1994), Mills (1999), Nicolau (1994, 1997), entre outros.

<sup>4</sup> As primeiras investigações alvo de publicações científicas nesta área devem-se a Merton (1980), Poterba e Summers (1986), French *et al.* (1987), Schwert (1989), Schwert (1990), Taylor e Xu (1997), entre outros. Contudo, estes autores não apresentam uma fundamentação teórica para a RV.

obtidos a partir de dados registados em intervalos fixo cada vez mais pequenos (dados de alta frequência intra-diários). A sua fundamentação teórica tem por base o conceito de semimartingala, teoria da Variação Quadrática (VQ), e a classe de modelos em tempo contínuo. O mecanismo é o seguinte: admitindo que o processo do retorno segue uma martingala especial, pode-se demonstrar que o processo de VQ (por definição) pode ser estimada consistentemente pela soma dos produtos cruzados dos retornos intra-diários, ou seja a RV. Por outras palavras, a RV converge em probabilidade para a VQ, quando se admite que o número de observações intra-diários tende para mais infinito ou quando intervalo entre as observações tende para zero (isto é, quando aumenta a frequência amostral). Aumentando a frequência amostral obtém-se um estimador arbitrariamente mais preciso da volatilidade. Sob determinadas condições do processo da média condicional a RV é um estimador não enviesado da matriz de variância covariância condicional. Restringindo à classe de modelos em tempo contínuo a VQ, sob determinadas condições reduz-se a Volatilidade Integrada (VI).

Apesar de recente, os desenvolvimentos teóricos e empíricos na literatura sobre RV são já bastante extensos. Um número substancial de trabalhos tem sido realizado com vista a investigar as propriedades da RV, nomeadamente nos seguintes aspectos: (i) estudo dos aspectos condicionais e não condicionais da distribuição da RV, no que se refere às características e regularidades empíricas normalmente observadas nas sucessões cronológicas financeiras; e (ii) estudo da sua performance na modelação e previsão da volatilidade, confrontados com os resultados obtidos com os modelos ARCH e de VE. A este propósito destacam-se muitos trabalhos, como por exemplo, os artigos de ABDL (2003) que investigam as taxas de câmbio (Marco/Dólar, Marco/Yen e Yen/Dólar),

Andersen, Bollerslev, Diebold e Ebens (2001) (doravante ABDE) que investigam os trinta activos subjacente do índice *Dow Jones*, Ebens (1999) que investiga o índice *Dow Jones*, Areal e Taylor (2002) que investigam o índice FTSE 100 (*stock index futures*), entre outros. As principais conclusões dos resultados obtidos são: (i) a distribuição do retorno estandardizado pelo desvio padrão (raiz quadrada da RV) é aproximadamente normal; (ii) a distribuição da RV não é normal e é fortemente enviesado à direita, enquanto que a distribuição do logaritmo da RV é aproximadamente normal; e (iii) uma das características importante desses estudos é a de que, o logaritmo da RV apresenta uma forte dependência temporal e parece ser bem descrita por um processo de memória longa.<sup>5</sup> Muitos autores usam o modelo ARFIMA (*Autoregressive Fractionally Integrated Moving Average*) com o objectivo de captar a dinâmica do processo. O resultado é de que estes modelos parecem descrever bem o comportamento da RV e proporcionam uma *performance* favorável quando comparado em termos de previsão com os modelos de volatilidade mais populares.

As possíveis motivações para uma profícua investigação nesta área estão relacionadas com dois aspectos relevantes. Primeiro, a crescente disponibilidade de dados registrados em intervalos de tempo cada vez mais curtos, juntamente com os avanços tecnológicos na área de computação, abriram espaço para a investigação empírica voltada para a análise e aplicação dos dados de alta-frequência em uma diversidade de aspectos dos mercados financeiros. O interesse no uso de dados de alta-frequência para o estudo da volatilidade em séries financeiras tem se revelado mais intensamente nos últimos anos. Segundo, as limitações apresentadas pelos procedimentos tradicionais mais populares.

---

<sup>5</sup> No caso multivariado a covariância não é normal, enquanto que a distribuição marginal das correlações é aproximadamente normal. As correlações exibe também efeito memória longa [ver por exemplo ABDE (2001)].

De um modo geral a maior parte das propostas na literatura para a modelação de séries financeiras são, em geral, deficientes para captarem a dinâmica dos processos. Os modelos ARCH e de VE multivariados são intratáveis do ponto de vista de estimação dado ao aumento de número de parâmetros a estimar.

Para além disto, o novo método oferece inúmeras vantagens: uma das vantagens da RV resulta principalmente da forma simples como é obtido e por ser uma variável observável. Ao contrário dos modelos da família ARCH e de VE, a RV não depende da especificação dos dois primeiros momentos condicionais, ou seja da forma funcional da média e da volatilidade, permitindo assim que a volatilidade seja modelada directamente. Uma outra vantagem do retorno intra-diário para estimar a volatilidade é a de que a precisão aumenta substancialmente (as investigações empíricas realizadas sobre a RV tem vindo a reforçar esta ideia). A possibilidade de se obter estimativas mais precisas de volatilidade tem despertado o interesse de muitos investigadores.

Uma questão muito problemática na construção da RV tem a ver com a escolha de frequência amostral, uma vez que uma das características dos dados de alta-frequência é a presença de *market microstructure effects*. Portanto, usar dados de frequência elevada para medir a volatilidade não é necessariamente a melhor aproximação já que a medida pode ser contaminada por *market microstructure effects*. Na prática, a escolha da frequência a utilizar depende do balanço (equilíbrio) entre a necessidade de uma amostra de alta-frequência para aproximar a modelos em tempo contínuo e o “custo” do

enviesamento devido ao *market microstructure effects*.<sup>6</sup> A solução geralmente adoptada na literatura consiste em considerar uma frequência intermédia. Os trabalhos empíricos situam normalmente a frequência entre 5 e 30 minutos.<sup>7</sup>

O objectivo da presente dissertação consiste em fazer uma aplicação empírica da RV ao índice PSI20, focando sobretudo na análise das propriedades das distribuições condicionais e não condicionais e confrontando com os resultados já obtidos na literatura. Assim, no segundo capítulo apresenta-se de uma forma breve o enquadramento teórico da RV. No terceiro analisam-se as propriedades das distribuições condicionais e não condicionais da RV do PSI20, e propõe-se um modelo dinâmico para RV. Finalmente, no quarto capítulo apresenta-se as principais conclusões.

---

<sup>6</sup> A ideia é de que quando se aumenta a frequência amostral o *market microstructure effects* (por exemplo, *bid/ask spreads* e *non-synchronous trading*) induz autocorrelação nos retornos intra diários, enviesando assim a RV. Esta questão tem sido muito investigada na literatura e não vai ser objecto de estudo no presente trabalho; ver, por exemplo, Andreou e Ghysels (2002), Campbell, Lo, MacKinlay (1997), Madhavan (2000), O'Hara (1995), Oomen (2002), Wood (2000), entre outros.

<sup>7</sup> Uma forma alternativa de eliminar a autocorrelação consiste no uso de filtros, usado por exemplo, nos trabalhos de Ebens (1999), ABIDE (2001), Thomakos e Wang (2003) (*moving average*) e Bollen e Inder (2002) (*autoregressive*), entre outros.

## 2. Breve Enquadramento Teórico da *Realized Volatility* (RV)

### 2.1. Variação Quadrática (VQ) e RV.

Considere-se  $p = \{p_t\}_{t \in [0, t]}$  o processo do logaritmo do preço em tempo contínuo de dimensão  $N$ , definido sobre um espaço de probabilidades  $(\Omega, F, P)$ , onde,  $\Omega$  é o conjunto de todos os estados da natureza (ou cenários de mercado),  $F$  é a álgebra- $\sigma$  de subconjuntos de  $\Omega$  e  $P$  é uma probabilidade sobre  $F$ . Seja  $\{F_t\}_{t \in [0, t]} \subseteq F$  uma filtração, isto é, uma família de sub-álgebras de  $F$  tal que  $F_s \subset F_t$  quando  $s \leq t$ . Vamos considerar a filtração natural  $F_t = \sigma(p_s; 0 \leq s \leq t)$  gerada pelo próprio processo até ao momento  $t$ , a qual contém toda a “história” do processo de preços dos activos (podendo incluir outras variáveis de estado).

Admitindo que o processo do retorno,  $r = \{r_t\}_{t \in [0, t]}$  [representemos o retorno (continuamente composta) no intervalo  $[t-h, t]$  por  $r_{t,h} = p_t - p_{t-h}$ ], não permite oportunidades de arbitragem e que tem média instantânea finita, tem-se que o processo de preço (e do retorno) pertence à classe de semimartingala especial, e tem uma decomposição canónica única – componente previsível de variação finita mais uma martingala local [ver Back (1991) e Protter (1992)]. Diz-se que  $p_t$  é uma semimartingala se poder ser decomposto como:

$$(1) \quad p_t = \alpha_t + m_t$$

onde  $\alpha_t$ , a componente *drift*, é um processo com trajectórias de variação finita localmente limitada e  $m_t$  é uma martingala local. Quando  $\alpha_t$  é um processo previsível<sup>8</sup> diz-se que é uma semimartingala especial.

Considerando que  $p_t$  pertence à classe de semimartingala especial, define-se o processo de VQ matricial  $N \times N$  para o processo do retorno associado como [ver, por exemplo, ABDL (2003) e Protter (1992)]:

$$(2) \quad [r, r]_t = p - \lim_{m \rightarrow \infty} \sum_{j=0}^{m-1} (r_{t_{j+1}} - r_{t_j})(r_{t_{j+1}} - r_{t_j}),$$

para qualquer decomposição  $t_0 = 0 < t_1 < \dots < t_m = t$  com  $\sup_j \{t_{j+1} - t_j\} \rightarrow 0$  para  $m \rightarrow \infty$ .

Pode-se assim concluir que a soma do produto cruzado dos retorno de alta-frequência intra-diários, a que se convencionou chamar de *Realized volatility* (RV), é um estimador consistente da VQ, dado por  $[r, r]_t$ . Por outras palavras, a RV converge em probabilidade para a VQ, quando se admite que o número de observações intra-diários tende para mais infinito ( $m \rightarrow \infty$ ) ou quando intervalo entre as observações (frequência amostral) tende para zero ( $\Delta = h/m \rightarrow 0$ )<sup>9</sup>.

ABDL (2003) demonstram uma outra forma de estabelecer relação entre a VQ e a volatilidade, considerando que a matriz de covariância condicional do retorno no momento  $t$ , onde  $0 \leq t \leq t+h$ , é dado pela soma do valor esperado (condicionado) do

<sup>8</sup> O valor no momento  $t$  de um processo previsível é conhecido um instante antes do momento  $t$ , como é o caso por exemplo de tendências determinísticas e todos os processos cãglãd (processos contínuos à esquerda e com limite à direita).

<sup>9</sup> Geralmente, considera-se intervalos igualmente espaçados:  $t_i = t_{i-1} + \Delta$ , onde  $\Delta = h/m$  e  $h = (t_m - t_0)$ .



incremento da VQ, que por sua vez é consistentemente estimado pela RV, mais uma componente aleatório (*erro*), ou seja

$$(3) \quad Cov[r_{t+h} - r_t | F_t] = E[[r, r]_{t+h} - [r, r]_t | F_t] + erro.$$

O erro é uma componente aleatória empiricamente relevante do desvio entre a RV e a matriz de covariância condicionada. Caso o  $erro \approx 0$  tem-se que  $Cov[r_{t+h} - r_t | F_t] \approx E[[r, r]_{t+h} - [r, r]_t | F_t]$ ; sob esta condição a RV será um estimador não enviesado da matriz de covariância condicionada por  $F_t$ . Pode-se provar [ver ABDL (2003)] que a componente aleatória pode ser nula na presença de vários efeitos, nomeadamente no caso da média ser constante. Até mesmo no caso em que a componente aleatória não é nula, pode demonstrar-se que é quanto muito de ordem  $h^{3/2}$ . Uma vez que a VQ é de ordem  $h$ , um  $h$  pequeno implica uma aproximação muito precisa. Nestas circunstâncias a RV, em contraste com os modelos tradicionais mais populares, proporciona um estimador não enviesado da matriz de covariância condicional, sem ser tomado qualquer modelo subjacente.

Uma vez que os resultados acima mencionados foram obtidos sem um modelo para volatilidade do retorno, vamos considerar uma classe de modelo definida em tempo contínuo, abordado no ponto seguinte.

## 2.2. Volatilidade Integrada (VI) e RV.

Considere  $p_t$  um processo de difusão governado pela equação diferencial estocástica multivariado

$$(4) \quad p_{t+h} - p_t = \int_0^h \mu_{t+\tau} d\tau + \int_0^h \sigma_{t+\tau} dW_\tau$$

onde  $\mu_t$  é uma função previsível integrada,  $W_t$  é o processo de Wiener Padrão de dimensão  $N$  e  $\sigma_t$  é o processo de difusão matricial  $N \times N$  definida positiva e estritamente estacionária. Trata-se de uma classe de processos muito popular que pertence a classe de semimartigala, muito usado na teoria financeira.

Admitindo que  $\mu_t$  e  $\Omega_t = \sigma_t \sigma_t'$  é independente de  $W_t$  no intervalo  $[t, t+h]$ , ABDL (2003) demonstram que

$$(5) \quad r_{t+h} | \sigma\{\mu_{t+\tau}, \Omega_{t+\tau}\}_{\tau=0}^h \sim N\left(r_t + \int_0^h \mu_{t+\tau} d\tau, \int_0^h \Omega_{t+\tau} d\tau\right),$$

onde  $\sigma\{\mu_{t+\tau}, \Omega_{t+\tau}\}_{\tau=0}^h$  é a álgebra- $\sigma$  gerada pela trajectória amostral de  $\mu_t$  e  $\Omega_t$  para  $0 \leq \tau \leq h$ . O incremento de VQ do retorno dado por

$$(6) \quad [r, r]_{t+h} - [r, r]_t = \int_0^h \Omega_{t+\tau} d\tau,$$

é normalmente designada por VI, que não sendo observável vai ser objecto de estimação. Note-se que tomando o valor esperado condicionado em (6) obtém-se a matriz de covariância condicionada definido em (3). Neste caso, da teoria da VQ definida em (1) e de (6), tem-se claramente que para qualquer  $t$ , a medida que a frequência amostral do retorno aumenta, ou  $\Delta \rightarrow 0$



$$(7) \quad \sum_{j=1, \dots, h[\Delta/h]} \left( r_{t+j\frac{\Delta}{m}} - r_{t+(j-1)\frac{\Delta}{m}} \right) \left( r_{t+j\frac{\Delta}{m}} - r_{t+(j-1)\frac{\Delta}{m}} \right)' - \int_0^h \Omega_{t+\tau} d\tau \rightarrow 0.$$

Observa-se algumas das características interessantes desta teoria:

- (i) a RV pode ser usada para aproximar as várias medidas *ex ante* de volatilidade – VQ, IV e matriz de covariância condicionada. Em contraste com os procedimentos tradicionais, a VI e a matriz de covariância condicionada do retorno podem ser obtidas através de um procedimento não paramétrico [a VQ e o seu estimador (RV) não dependem da especificação da média e nem da volatilidade], pela soma do quadrado e do produto cruzado do retorno de alta-frequência intra-diário;
- (ii) a RV é directamente observável, em contraste com a sua contraparte teoria subjacente, o processo de VQ;
- (iii) para  $m$  suficientemente grande, porém, a RV proporciona uma aproximação arbitrariamente boa para a VQ;
- (iv) a convergência é obtida sem envolver qualquer tipo de estabilidade estrutural (estacionaridade).

### 3. Aplicação Empírica ao índice PSI20

Neste capítulo iremos fazer uma aplicação empírica da RV, apresentando no capítulo anterior. O objectivo desta aplicação é investigar as propriedades das distribuições condicionais e não condicionais da RV, sobretudo, e modelar as suas principais características dinâmicas.

#### 3.1. Dados e cálculo da RV

A nossa análise empírica da RV incide sobre as cotações/observações intra-diários do índice PSI20 da *Euronext Lisbon*, registrados no período 1/Janeiro/1998 a 22/Fevereiro/2004 (total de 1.519 observações diários) ver a figura 1.<sup>10</sup> O valor do índice é calculado em cada instante que houver transacções isto é, em intervalo irregulares. Neste caso, a escolha do intervalo entre as observações para o cálculo da RV deve ser pequeno o bastante para dar uma ideia precisa da volatilidade, mas não tão pequeno que gere *microstructure effects*, o que enviesaria a RV. Na prática a selecção de retornos em intervalos de tempo fixos depende do balanço (equilíbrio) entre a necessidade de uma amostra de alta-frequência para aproximar a modelos em tempo contínuo e o “custo” da *market microstructure biases*. Os intervalos entre 5 e 45 minutos têm sido uma prática comum [ver, por exemplo, Omen (2002), ABDL (2003), Ebens (1999), entre outros], para a qual as propriedades da série não seriam seriamente distorcidas por aspectos de *market microstructure effects*. No nosso estudo vamos considerar duas frequências: 5 e 30 minutos (designamos de séries I e II, respectivamente).

---

<sup>10</sup> O autor agradece a *Euronext Lisbon* pela disponibilização dos dados.

Considerando o logaritmo do índice PSI20 geral ( $p_t$ ), calcula-se o retorno diário pela soma dos retornos intra-diários – ver a figura 2.<sup>11</sup> Para ambas as séries obtém-se a RV diário,  $\sigma_t^2$  ( $h = 1$ ), pela soma dos quadrados dos retornos intra-diários, representado na figura 3.

Figura 1: Índice PSI20 (1/Janeiro/1998 a 22/Fevereiro/2004)

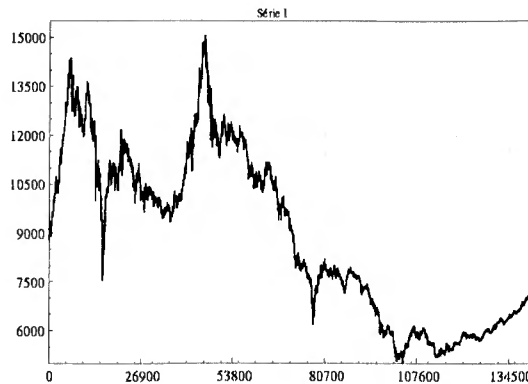
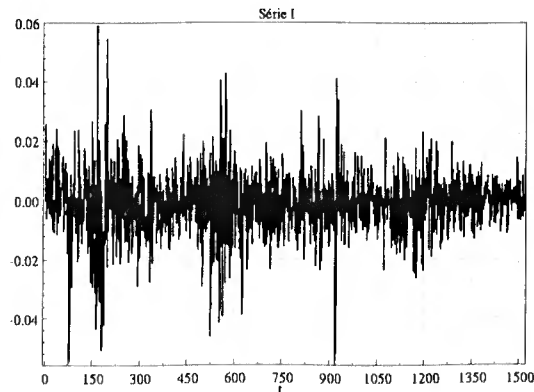
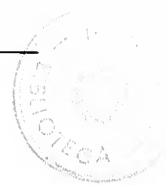
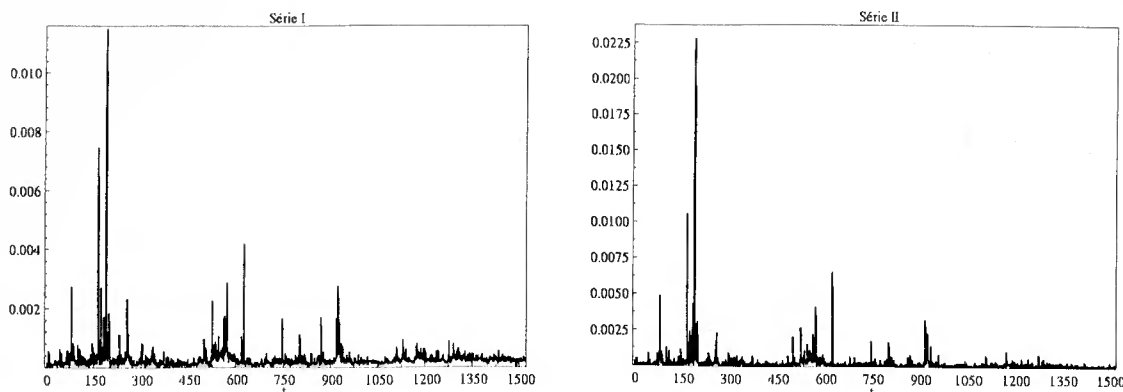


Figura 2: Retorno Diário do Índice PSI20 ( $r_t$ )



<sup>11</sup> Com o objectivo de homogeneizar a distribuição das observações, optou-se por considerar apenas as observações verificadas das 9 às 16 horas. O valor do índice calculado equidistante baseia-se no método *previous tick*, proposto por Wasserfallen e Zimmermann (1985): para qualquer ponto do tempo,  $\tau \in [t_0, t_n)$   $PSI20(\tau) \equiv PSI20(t_j)$ ,  $\tau \in [t_j, t_{j+1})$ .

Figura 3: *Realized Volatility* ( $\sigma_t^2$ )

É notória através da representação gráfica de  $\sigma_t^2$  para ambas as séries (ver a figuras 3), períodos de volatilidade mais elevada intercalados com períodos de menor volatilidade [grandes (pequenas) variações tendem a ser seguidas por grandes (pequenas) variações, positivas ou negativas] – fenómeno típico das séries financeiras que já é conhecido na literatura por *volatility clustering* [Mandelbrot (1963) e Fama (1965)]<sup>12</sup> – o que indica que a volatilidade não pode ser constante ao longo do tempo. A implicação da *volatility clustering* é de que um choque na volatilidade hoje influencia a expectativa da volatilidade nos períodos futuros. Esta característica está directamente em linha com os resultados obtidos na literatura dos modelos ARCH e de VE e RV [ver Kim e Kon (1994), ABDE (2001) e ABDL (2003)]. É notória uma diminuição da volatilidade (RV) para o fim do período. Verifica-se ainda que a RV construída a partir da série I apresenta uma menor variabilidade do que aquela construída a partir da série II.

<sup>12</sup> Ver, por exemplo, Baillie *et al.* (1996), Chou (1988) e Schwert (1989).

### 3.2. Propriedades das Distribuições não Condicionais do Retorno e da RV

Neste ponto iremos identificar algumas propriedades estatísticas não condicional do retorno, retorno estandardizado, RV e o logaritmo da RV, principalmente no que se refere ao grau de achatamento e assimetria.

#### 3.2.1. Retorno e Retorno Estandardizado

Na tabela 1 abaixo apresentada encontra-se sumariado algumas estatísticas [média, variância, *Kurtosis*, *Skewness*, máximo, mínimo e o *p-value* da estatística teste JB [Jarque e Bera (1987) ] da distribuição empírica do retorno diário ( $r_t$ ) e do retorno diário estandardizado pelo desvio padrão ( $r_t/\sigma_t$ ) considerando as duas séries.

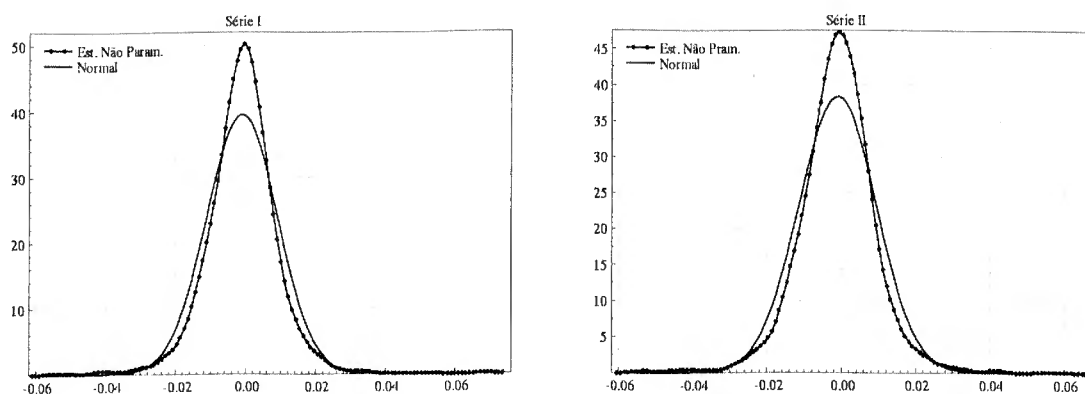
Tabela 1: Distribuição não Condicional do Retorno e Retorno Estandardizado

	Série	Média	Variância	<i>Kurtosis</i>	<i>Skewness</i>	Mínimo	Máximo	<i>p-value</i> JB
$r_t$	I	-1.2e-3	1.1e-4	7.35	-0.14	-0.06	0.06	2.3e-262
	II	-1.2e-3	1.0e-4	8.26	-0.05	-0.06	0.07	0.00
$r_t/\sigma_t$	I	-0.06	0.41	3.97	0.26	-2.34	2.83	2.0e-17
	II	-0.08	0.58	3.94	0.18	-2.84	3.25	1.1e-14

De acordo com os valores apresentados na tabela, podem-se tirar as seguintes conclusões relativamente à distribuição empírica de  $r_t$ : (i) apresenta um grau de achatamento ou *Kurtosis* superior à distribuição normal (3), o quer dizer que a distribuição empírica do retorno é leptocúrtica. As distribuições leptocúrticas indicam basicamente que ocorrem com frequência superior ao que seria de esperar se os dados

tivessem sido gerados por distribuições normais, valores muito altos ou muito baixos. Por outras palavras, é indicativo que o mercado atribuiu grande probabilidade à ocorrência de acontecimentos extremos; (ii) tem o coeficiente de assimetria ou *Skewness* negativo (i.e. enviesamento à esquerda) inferior ao valor da distribuição normal (0), o que evidencia que o mercado atribuiu maior probabilidade a fortes descidas no valor do índice do que a fortes subidas [ou seja, o número de valores muito baixos (negativos) foi superior ao número dos valores muito altos]; (iii) O *p-value* da estatística teste JB (praticamente nula) confirma claramente que o retorno diário não é normal, resultado consistente com os resultados já obtidos na literatura. Na figura 4 é representada graficamente a distribuição de  $r_t$  (estimativa da densidade de Kernel)<sup>13</sup>, juntamente com a distribuição normal de média e variância igual à distribuição empírica. A figura mostra que a distribuição é leptocúrtica, embora não seja claro a partir dela, que seja assimétrica negativa.

Figura 4: Estimativa da Densidade do retorno ( $r_t$ ) versus Densidade Gaussiana

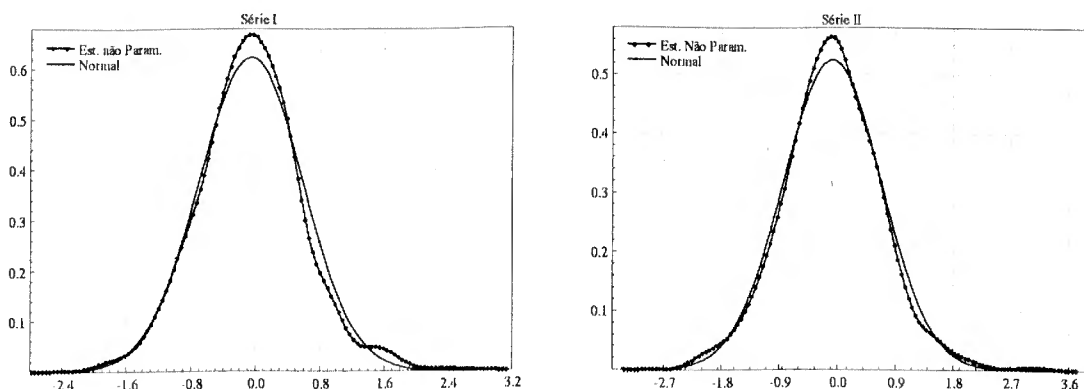


<sup>13</sup> Estimativa não paramétrica da densidade de  $r_t$  utilizando-se um kernel Gaussiano e uma *bandwidth*  $h = 1.90 \times \hat{\sigma}r_t n^{-1/5}$  onde  $\sigma r_t$  é o desvio padrão de  $r_t$  [ver Silverman (1986)]. Representações gráficas semelhantes são apresentadas ao longo do trabalho.



Quanto à distribuição empírica de  $r_t/\sigma_t$ , em contraste com distribuição empírica de  $r_t$ , ambas as séries apresentam uma distribuição empírica aproximadamente normal. Em particular, a *Kurtosis* e *Skewness* [(3,97; 0,26) e (3,94; 0,18) das séries I e II, respectivamente], tem valores ligeiramente superiores aos da distribuição normal. Este é também evidente através da figura 5, onde é representada graficamente a estimativa não paramétrica da densidade de kernel Gaussiano de  $r_t/\sigma_t$ , juntamente com a referência da densidade normal. No entanto, o teste BJ rejeita a normalidade, mesmo para níveis de significância relativamente reduzidos. Este resultado é consistente com os obtidos por ABDE (2001) e ABDL (2003). A distribuição dos retornos estandardizados pela raiz quadrada da RV tende a ser menos leptocúrtica do que a distribuição dos retornos estandardizados pela estimativa ARCH [ver, por exemplo, Bollerslev *et al.*, (1994) e Andersen e Bollerslev (1998)].

Figura 5: Estimativa da Densidade de  $r_t/\sigma_t$  versus Densidade Gaussiana



### 3.2.2. RV

A tabela 2 documenta também o mesmo conjunto do sumário estatístico apresentado no quadro anterior, para a distribuição não condicional da RV e o do logaritmo da RV. Ao

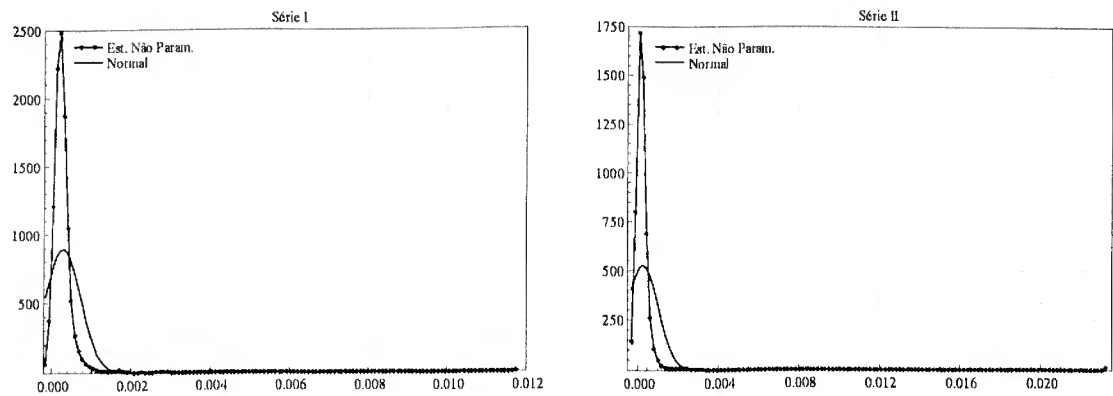
observar os valores apresentados na tabela, resulta claro que (para ambas as séries)  $\sigma_t^2$  não é normal, enquanto que o  $l\sigma_t^2$  apresenta uma distribuição aproximadamente normal.

Tabela 2: Distribuição não Condicional de  $\sigma_t^2$  e de  $l\sigma_t^2$

	Série	Média	Variância	Kurtosis	Skewness	Mínimo	Máximo	p-value JB
$\sigma_t^2$	I	2.8e-4	2.0e-7	320.07	14.87	2.7e-5	0.01	0.0
	II	2.4e-4	5.7e-7	557.13	20.44	7.9e-6	0.02	0.0
$l\sigma_t^2$	I	-8.49	0.44	5.55	0.80	-10,52	-4.46	2.6e-125
	II	-8.96	0.89	4.58	0.73	-11,75	-3.78	9.5e-65

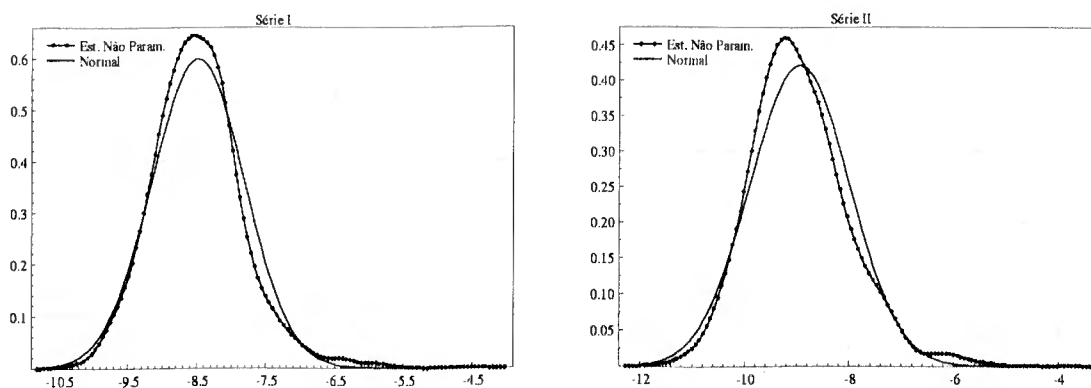
A distribuição empírica de  $\sigma_t^2$  (para ambas as séries) é fortemente leptocúrtica (tem “cauda pesada”) e fortemente enviesado à direita (assimétrica positiva). O valor da *Kurtosis* e *Skewness* [(320.07; 14.87) e (557.13; 20.44) séries I e II, respectivamente] é muito superior aos da distribuição normal (3; 0). O *p-value* da estatística de JB indica a rejeição da hipótese nula de normalidade a qualquer nível de significância [ver figura 6 - estimativa da densidade de Kernel e a distribuição normal de média e variância igual à distribuição empírica].

Figura 6: Estimativa da Densidade de  $\sigma_t^2$  versus Densidade Gaussiana



Tomando o logaritmo da RV ( $\ln \sigma_t^2$ ) os valores de *Kurtosis* e *Skewness* reduzem-se substancialmente, aproximando-se dos valores da distribuição normal. Em particular, a *Kurtosis* e *Skewness* [(5.55; 0,80) e (4.58; 0.76) das séries I e II, respectivamente] são inferiores comparativamente os valores obtidos com o  $\sigma_t^2$  [(320.07; 14.87) e (557.13; 20.44) séries I e II, respectivamente]. Resulta que a distribuição de  $\ln \sigma_t^2$  e  $\sigma_t^2$  da série II é mais leptocúrtica e mais assimétrica do que a série I. Estes resultados são evidentes através da figura 7. Embora seja aproximadamente normal, o *p-value* das estatísticas de JB indica a rejeição da hipótese nula de normalidade. Esta evidência é consistente com os resultados obtidos na literatura sobre a RV.

Figura 7: Estimativa da Densidade de  $\ln \sigma_t^2$  versus Densidade Gaussiana



É interessante, de facto, observar o seguinte: (i) a variância de  $\sigma_t^2$  e  $\ln \sigma_t^2$  da série I ( $2.0e-7$  e 0.44, respectivamente) é relativamente inferior à da Série II ( $5.7e-7$  e 0.89, respectivamente), traduzindo-se assim num aumento da precisão (ou seja, considerando uma maior frequência amostral, aumenta a precisão do estimador); (ii) ABDL (2000) argumentam que, se *microstructure effects* está presente a RV pode diferir de acordo com a frequência amostral. Como pode observar pela tabela 2 a média de  $\sigma_t^2$  e  $\ln \sigma_t^2$  das duas séries é estável.

### 3.3. Propriedades das Distribuições Condicionais da RV

A distribuição condicional da volatilidade dos activos financeiros tem sido objecto de uma ampla investigação nas últimas décadas. Motivado pelos resultados da análise precedente, neste ponto investigaremos as propriedades das distribuições condicionais de  $\sigma_t^2$ , onde focaremos a nossa atenção no estudo da dependência temporal (memória longa) da volatilidade e no efeito assimétrico entre a volatilidade e os retornos (o retorno e a volatilidade são negativamente correlacionados; choques positivos e negativos produzem efeitos diferentes na volatilidade).

#### 3.3.1. Dependência Temporal e Memória Longa

Uma das regularidades empíricas do retorno dos activos financeiros é a dependência temporal da volatilidade. A literatura sobre a modelação da volatilidade tem documentado que tal dependência é fortemente persistente, como se pode constatar pela figura 8 abaixo apresentado. A persistência da volatilidade é mais notória através do espectro da figura 9, onde é representada graficamente a função de autocorrelação (FAC) de  $\sigma_t^2$  da série I (a série II apresenta um comportamento muito semelhante) até o desfasamento (*lag*) de ordem 100. É visível que a autocorrelação mantém-se significativa para *lags* relativamente alto, está sistematicamente acima do intervalo de confiança, e tem um decaimento lento (hiperbólico) para zero (ver os valores das autocorrelações na tabela 3) – características típicas dos processos de memória longa [comportamento semelhantes foram encontradas em ABDE (2001), ABDL (2003), Ebens (1999), Areal e Taylor (2002), entre outros].

Figura 8: Logaritmo da *Realized Volatility* ( $\ln \sigma_t^2$ )

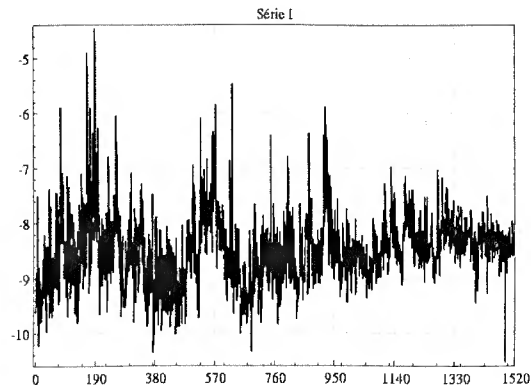


Figura 9: FAC de  $\ln \sigma_t^2$

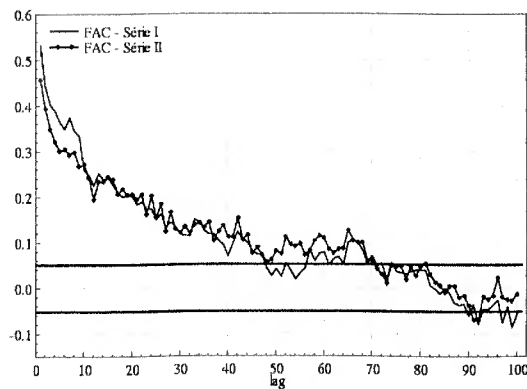
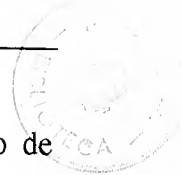


Tabela 3: Dependência dinâmica de  $\ln \sigma_t^2$

					Autocorrelação				
	Série	Pvalue LB	ADF <sup>(*)</sup>	d					
					lag 1	lag 5	lag 10	lag 15	lag 20
$\ln \sigma_t^2$	I	0.00	-5.08	0.31	0.53	0.35	0.23	0.20	0.17
	II	0.00	-4.63	0.29	0.46	0.30	0.20	0.22	0.20

(\*) Teste Dickey-Fuller Aumentado com drift; valor crítico 5% = -2.87.

Na tabela 3, é também apresentado o *p-value da* estatística teste de LB [Ljung e Box (1978)], o teste Dickey-Fuller Aumentado (DFA) [Dickey e Fuller (1981)] e ordem de



integração fraccionaria  $d$  [obtido através do estimador semiparamétrico Gaussiano de Robinson (1995)]. Analisando o  $p$ -value da estatística teste de LB para a significância conjunta das primeiras 22 correlações de  $\sigma_t^2$ , a hipótese de autocorrelação nula (ruído branco) é claramente rejeitada (confirmando o resultado visual do correlograma). O valor reduzido das autocorrelações, junto com o decaimento lento, sugere que  $\sigma_t^2$  não contém uma raiz unitária, mas a presença de memória longa. Como se pode constatar o teste DFA para a constante e 22 lags, rejeita a hipótese nula de raiz unitária,  $I(1)$ , o que sugere que  $\sigma_t^2$  seja integrada fraccionária. A partir do valor do parâmetro da ordem de integração fraccionaria  $d$  (0.31 e 0.29 para as série I e II, respectivamente), conclui-se que em ambas as séries  $\sigma_t^2$  é um processo fraccionalmente integrado de covariância estacionária (tem a propriedade de memória longa e é estacionária).<sup>14</sup> Os resultados obtidos na literatura sobre RV é de que o  $\sigma_t^2$  é um processo fraccionalmente integrado de covariância estacionária e o valor de  $d$  situam-se à volta dos 0.4.

### 3.3.2. Efeito Assimétrico

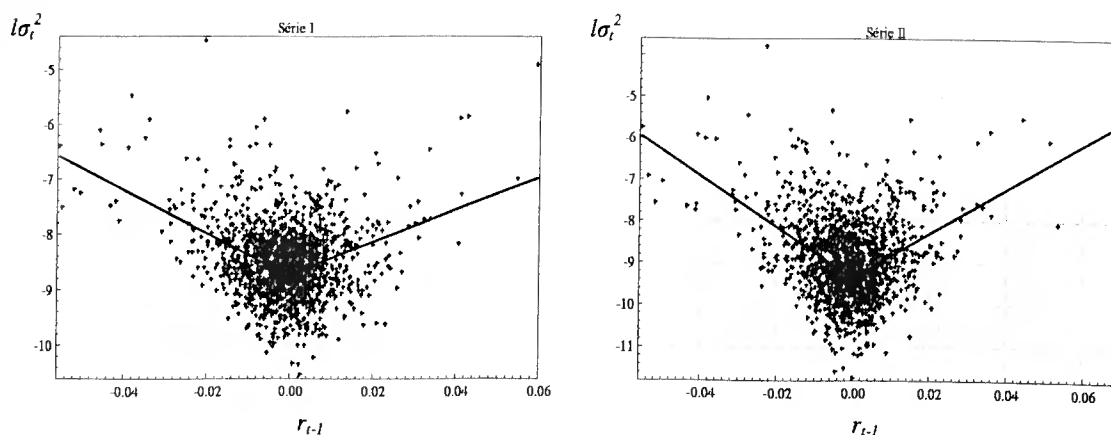
Nos mercados financeiros, observa-se frequentemente uma relação assimétrica entre a volatilidade e os retornos, no sentido de que, a volatilidade tende a aumentar mais com o aparecimento de uma *bad news* do que com uma *good news*. Geralmente, a volatilidade é maior depois de choques negativos do que depois de choques positivos de mesma magnitude.<sup>15</sup> Estudos prévios têm documentado esta relação assimétrica entre a volatilidade e os retornos; ver, por exemplo, Black (1976), Pagan e Schwert (1990) e

<sup>14</sup> Um processo tem a propriedade de memória longa se  $0 < d < 1$  e tem memória curta se  $-1 < d < 0$  e; Se  $0 < d < 0.5$ , o processo é estacionário, tem memória longa e diz-se que é persistente; se  $-0.5 < d < 0$ , o processo tem memória curta e diz-se que é antipersistente.

<sup>15</sup> Este comportamento assimétrico da volatilidade está na origem dos modelos GARCH exponencial [EGARCH, Nelson (1991)] e Threshold ARCH [TARCH, Zakoian (1990, 1994)].

Engle e Ng (1993). A figura 10 abaixo apresentada sugere uma relação assimétrica entre a volatilidade e o retorno desfasado para as duas séries.

Figura 10: Assimetria de  $l\sigma_t^2$



Para investigar assimetria na influência de retornos passados negativos e positivos na volatilidade do índice, estimamos o seguinte modelo usando método de mínimos quadrados

$$l\sigma_t^2 = w + \gamma r_{t-1} + \phi r_{t-1} I(r_{t-1} < 0) + u_t$$

onde o indicador  $I(r_{t-1} < 0)$  toma valor 1 se  $r_{t-1} < 0$ , e zero caso contrario. Na tabela 4 são apresentados os resultados da estimação do modelo com todos os parâmetros estatisticamente significativo. Verifica-se que o coeficiente  $\phi$  é negativo  $[(-64.59, -101.35)$  das série I e II, respectivamente] e superior (em valor absoluto) ao coeficiente do retorno  $\gamma$  (27.19 e 44.06 para as série I e II, respectivamente) indicando a existência do efeito assimétrico. O mesmo resultado chegou Ebens (1999).

Tabela 4: Assimetria do de  $l\sigma_t^2$  (\*)

	Série	$w$	$\gamma$	$\varphi$	$R^2$
$l\sigma_t^2$	I	-8.73 (0.02)	27.19 (2.89)	-64.59 (4.38)	0.14
	II	-9.33 (0.03)	44.06 (4.13)	-101.35 (6.27)	0.16

(\*) Os valores entre parênteses representam o desvio padrão.

### 3.4. Modelação da RV

Em geral, os modelos estatísticos desenvolvidos para estimação, modelação e previsão da volatilidade tentam reproduzir um conjunto de características e regularidades empíricas normalmente observadas em séries financeiras. Deste modo, nesta parte do trabalho propomos modelar o logaritmo da RV, considerando os principais efeitos detectados no ponto anterior (memória longa e assimetria), e depois verificamos se o modelo estimado modela convenientemente as principais características observadas. Tratando a volatilidade como uma variável observada em vez de latente permite usar técnicas de série temporal empregues quando se modela a média condicional, contrariamente aos modelos ARCH e VE onde a volatilidade é modelada e estimada simultaneamente.

A análise precedente mostrou que o  $l\sigma_t^2$  é não constante ao longo do tempo e, além disso, é condicionalmente dependente da informação passada e fraccionalmente integrado de covariância estacionária. Assim, parece correcto assumir-se modelo que o  $l\sigma_t^2$  pode ser bem modelado por um modelo ARFIMA, embora outros modelos possam



também ser considerados.<sup>16</sup> Neste contexto, tendo em conta a presença de memória longa e a correlação negativa entre a volatilidade o retorno, propomos modelar o  $\sigma_t^2$  usando o seguinte modelo ARFIMAX( $p, d, q$ ):

$$\alpha(L)(1-L)^d(l\sigma_t^2 - \mu_t) = \beta(L)\varepsilon_t, \quad \mu_t = \omega + \gamma r_{t-1} + \varphi r_{t-1}I(r_{t-1} < 0)$$

onde,  $d$  representa a ordem de integração fraccionária,  $\alpha(L)$  e  $\beta(L)$  representam os polinómio da componente autoregressiva e média móvel de grau  $p$  e  $q$ , respectivamente, sem zeros no círculo unitário e o  $\varepsilon_t$  é processo ruído branco de media zero com variância  $\varepsilon_t^2$ . O  $\mu_t$  é o valor esperado de  $l\sigma_t^2$  (o qual inclui o efeito assimétrico).

Na tabela 5 abaixo representmos os resultados da estimação com todos os parâmetros estatisticamente significativos. Na tabela é dado também o *Schwartz Bayesian Information Criterion* (SBC), a variância residual ( $\sigma_{\text{residual}}^2$ ) e o *p-value* da estatística teste LB para as primeiras 38 correlações residual.<sup>17</sup> Pela análise da tabela podemos tirar as seguintes conclusões: primeiro, verifica-se que o valor estimado do parâmetro  $d$  para o modelo ARFIMAX [(0.43, 0.37) série I e II, respectivamente] garante a estacionaridade e integração fraccionária, e são ligeiramente superiores ao do estimador de Robinson (1995) obtido anteriormente; segundo, tendo em atenção o valor estimado dos parâmetros  $\gamma$  e  $\varphi$  (de ambas as séries) indicam uma relação significativa entre  $\sigma_t^2$  e  $r_{t-1}$ ; terceiro, constata-se que a série I apresenta uma menor variância residual da série II

<sup>16</sup> Muitos estudos argumentam que a volatilidade dos mercados financeiros pode ser descrita por um processo de memória longa (modelos ARFIMA, modelos ARCH fraccionada integrada e de VE). [ver, por exemplo, Ding *et al.* (1993), Breidt *et al.* (1998), Robinson e Zaffaroni (1998), ABDE (2001), ABDL (2003), entre outros].

<sup>17</sup> Os parâmetros do modelo ARFIMA foram estimados usando o programa estatístico OX [Ver Ooms e Doornik (1999) para a documentação sobre ARFIMA *package*].

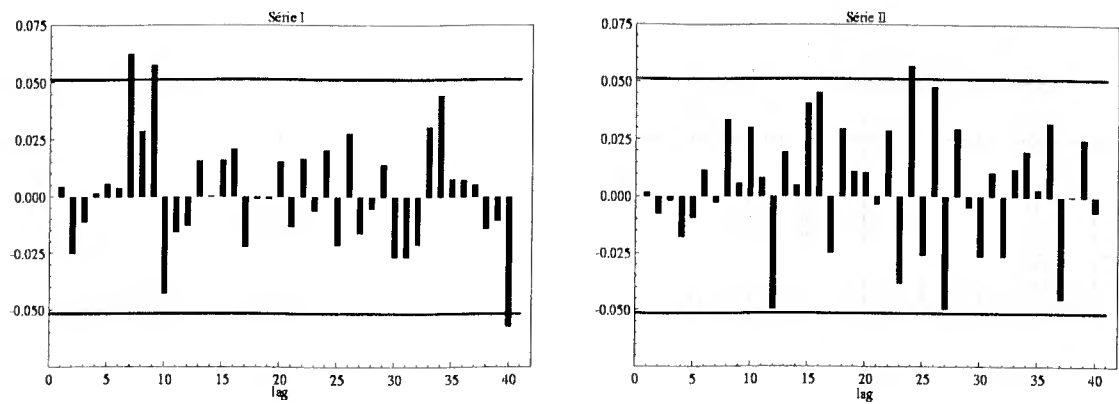
de menor frequência amostral; finalmente, o *p-value* da estatística teste de LB não rejeita a hipótese de autocorrelação nula, o que permite nos concluir que o modelo ARFIMA proposto descreve razoavelmente bem o  $\log \sigma_t^2$ , como se pode constatar através da FAC residual apresentado na a figura 10.

Tabela 5: Estimativas do Modelo ARFIMAX (\*)

Série	<i>d</i>	$\beta_1$	$\omega$	$\gamma$	$\varphi$	<i>SBC</i>	$\sigma^2_{residual}$	<i>p-value</i> LB <sub>38</sub>
ARFIMAX	I	0.43 (0.04)	-0.19 (0.05)	-8.64 (0.40)	13.49 (2.45)	-34.83 (3.90)	1.56	0.52
	II	0.37 (0.04)	-0.17 (0.05)	-9.15 (0.32)	17.55 (3.77)	-44.48 (5.98)	2.36	0.78

(\*) Os valores entre parênteses representam o desvio padrão. Os *p-value* da estatística *t* são todos praticamente nulos.

Figura 11: FAC Residual



## 4. Conclusões

O objectivo da dissertação consistiu em fazer uma aplicação empírica o novo método (não paramétrica) de estimação da volatilidade, recentemente proposto na literatura, a *Realized Volatility* (RV), obtida pela soma dos produtos cruzados dos retornos de alta-frequência intra-diários. Na introdução focámos o interesse do tema desenvolvido neste trabalho e os resultados já obtidos. No capítulo II fizemos um breve enquadramento teórico do método, que tem por base três elementos na sua fundamentação teórica: a teoria da Variação Quadrática (VQ), o conceito de semimartingala e a classe de modelos em tempo contínuo.

No capítulo 3 considerámos uma aplicação empírica da RV ao índice PSI20. O objectivo principal foi o de investigar as propriedades das distribuições condicionais e não condicionais da RV. Considerando duas séries com frequência amostral de 5 e 30 minutos verificámos para ambas que as distribuições empíricas da RV não são normal e são fortemente enviesado à direita, enquanto que as distribuições marginais do logaritmo da RV são aproximadamente normais, assim como são as distribuições do retorno estandardizado pelo desvio padrão (raiz quadrada da RV). O logaritmo da RV apresenta uma forte dependência temporal e parece ser bem descrita por um processo de memória longa. Os nossos resultados são consistentes com os resultados já obtidos na literatura. Verificamos ainda que logaritmo da RV exhibe efeito assimétrico da volatilidade.

Considerando a volatilidade como uma variável observável, em vez de latente como nos modelos ARCH e de VE, propusemos modelar as características dinâmicas do

logaritmo da RV através do modelo ARFIMA. Constatamos que o modelo ARFIMAX proposto modela razoavelmente bem o logaritmo da RV. Esse facto tem sido confirmado em muitos estudos.

Podemos focar alguns tópicos de interesse que poderão complementar o presente trabalho. Haverá interesse, num trabalho futuro, investigar se o modelo desenvolvido se pode adequar bem na modelação de outros fenómenos, como por exemplo, “alterações de regime”, efeitos dos dias da semana, efeitos de reversão, entre outros. Uma outra questão de grande interesse será ainda um estudo comparativo em termos de previsão com os modelos tradicionais mais populares. Seria também interessante fazer um estudo multivariado considerando os 20 activos que compõem o índice PSI20, assim como considerar índices de outros mercados financeiros.

## Bibliografia

- Andersen, T. G., Bollerslev T., (1998), Answering the skeptics: yes, standard volatility models do provide accurate forecasts, *International Economic Review* 39, 885-905.
- Andersen, T. G., Bollerslev T., Diebold F. X., Ebens H., (2001), The distribution of realized stock return volatility, *Journal of Financial Economics*, 61, 43-76.
- Andersen, T. G., Bollerslev T., Diebold F. X., Labys P., (2000), Market Microstructure Effects and the Estimation of Integrated Volatility, Work in progress, Duke University and University of Pennsylvania.
- Andersen, T. G., Bollerslev T., Diebold F. X., Labys P., (2003), Modeling and forecasting realized volatility, *Econometric*, 71, 579-625.
- Andreou, E., Ghysels, E. (2002), Rolling-sample volatility estimators: Some new theoretical, simulation, and empirical results, *Journal of Business and Economic Statistics*, 20(3), 363-376.
- Areal, N., Taylor, S.J., (2002), The realized volatility of FTSE-100 futures prices, *Journal of Futures Markets*, 22, 627-648.
- Back, K. (1991), Asset pricing for general processes, *Journal of Mathematical Economics*, 20, 371-395.
- Baillie R. T., Bollerslev T., Mikkelsen H. O., (1996), Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 74, 3 – 30.
- Barndorff-Nielsen, O.E., Shephard, N., (2001), Non-Gaussian OU based models and some of their uses in financial economics, *Journal of the Royal Statistical Society, Series B* 63, 167-241.
- Barndorff-Nielsen, O.E., Shephard, N., (2002), Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B* 64, 253-280.
- Bera, A. K., Higgins, M. L., (1993), ARCH models: Properties, estimation and testing, *Journal of Economic Surveys*, 7, 305-366.
- Black, F. (1976), Studies of Stock Market Volatility Changes, 1976 Proceedings of the American Statistical Association, *Business and Economic Statistics Section*, pp. 177-181.
- Bollen, B., Inder, B., (2002), Estimating daily volatility in financial markets utilizing intraday data', *Journal of Empirical Finance*, 9, 551-562.
- Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T., Engle, R. F., Nelson, D. B., (1994), *ARCH models*, In Engle, R. F., and McFadden, D. L.(eds.), *Handbook of Econometrics*, Vol. 4, Ch. 49. Amsterdam: North-Holland.
- Breidt, F.J., Crato, N., de Lima, P., (1998), On the detection and estimation of long-memory in stochastic volatility, *Journal of Econometrics*, 83, 325 – 348.

- Campbell, J. Y., A. W. Lo, A. C. MacKinlay, (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.
- Chou, R. Y. (1988), Volatility persistence and stock valuations: some empirical evidence using GARCH, *Journal of Applied Econometrics*, 3, 279 – 94
- Clark, P. K. (1973), A Subordinated Stochastic Process Model with finite Variance for Speculative Prices, *Econometrica*, 41:135-155.
- Dickey, D. A., Fuller, W. A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49, 1057-1072.
- Ebens, H. (1999), Realized stock volatility, Working paper 420, Johns Hopkins University.
- Engle R.F., (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987-1008.
- Engle, R. F., Ng. V. K. (1993), Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, 48, 1749-1778.
- Engle, R.F., Lilien, D. M.; Robinson, R.P., (1987), Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model, *Econometrica*, 55, n° 2, 391-407.
- Fama, E. (1965), The Behavior of Stock Market Prices, *Journal of Business*, 38:34-105.
- French, K.R., Schwert, G.W., Stambaugh, R.F., (1987), Expected stock returns and volatility, *Journal of Financial Economics*, 19, 3 – 29.
- Gourieroux, C. (1997), *ARCH Models and Financial Applications*. New York: Springer Verlag.
- Hamilton, J. D. (1994). *Time series Analysis*. Princeton: Princeton University Press.
- Hull, J., A. White (1987), The Pricing of Options on Assets with Stochastic Volatilities, *The Journal of Finance*, Vol XLII, 281-300.
- Jarque, C. M., Bera, A. K. (1987), A test for normality of observations and regression residuals, *International Statistical Review*, 55, 163--172.
- Kim, D., Kon, S.J., (1994), Alternative models for the conditional heteroscedasticity of stock returns, *Journal of Business*, 67, 563– 598.
- Ljung, G. M., Box, G. E. P. (1978), On a measure of lack of fit in time series models, *Biometrika*, 65, 297-303.
- Madhavan, A. N., (2000), Market Microstructure: A Survey, *Journal of Financial Markets*, 3, 205–258.
- Mandelbrot, B. B. (1963), The Variation of Certain Speculative Prices, *Journal of Business* 36:394-419.
- Merton, R. C. (1973), An Intertemporal Capital Asset Pricing Model, *Econometrica* 42:867-887.

- Merton, R.C. (1980), On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics*, 8, 323-361.
- Mills, T. C. (1999), *The Econometric Modelling of Financial Time Series*, Cambridge: Cambridge University Press.
- Murteira B.J.F., Müller D.A., Turkmann K.F., (1993), *Análise de Sucessões Cronológicas*, McGraw-Hill.
- Nelson, D.B., (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59, nº 2, 347-370.
- Nicolau J. C., (1997b), Definição, Identificação e Estimação do Modelo ARCH e Comparação com Outros Modelos de Volatilidade, *Estudos de Economia*, Vol. XVI-XVII, 393-409.
- Nicolau J.C., (1994), Modelos ARCH, Tese de Mestrado, ISEG/UTL, Lisboa.
- O'Hara, M., (1995), *Market Microstructure Theory*, Blackwell Publishers Ltd, Oxford, UK.
- Oomen, R. A. C. (2002), Modelling realized variance when returns are serially correlated, manuscript, Warwick Business School, The University of Warwick.
- Ooms, M., Doornik, J. A. (1999), Inference and forecasting for fractional autoregressive integrated moving average models, with an application to US and UK inflation, *Econometric Institute Report*, EI-9947/A, Erasmus University Rotterdam.
- Pagan, A. R., Schwert G. W., (1990), Alternative Models for Conditional Stock Volatility, *Journal of Econometrics*, 45:267-290.
- Poterba, J., Summers L., (1986), The Persistence of Volatility and Stock Market Fluctuations, *American Economic Review*, 76, 1124-1141.
- Protter, P. (1992), *Stochastic Integration and Differential Equations*, New York: Springer-Verlag.
- Robinson, P.M., (1995), Gaussian semiparametric estimation of long range dependence, *Annals of Statistics*, 23, 1630-1661.
- Robinson, P.M., Zaffaroni, P., (1998), Nonlinear time series with long memory: a model for stochastic volatility, *Journal of Statistical Planning and Inference*, 68, 359- 371.
- Schwert, G.W. (1989), Why Does Stock Market Volatility Change Over Time?, *Journal of Finance*, 44, 1115-1153.
- Schwert, G.W. (1990), Stock Volatility and the Crash of 87, *Review of Financial Studies*, 3, 77-102.
- Shephard, N. (1996), *Statistical aspects of ARCH and stochastic volatility*, In Cox, Hinkley D. R., e Barndorff-Nielsen D. V., O. E.(eds.), *Time Series Models in Econometrics, Finance and Other Fields*, pp. 1-67. London: Chapman & Hall.
- Silverman B.W., (1986), *Density Estimation For Statistics and Data Analysis*, Monographs on Statistics and Applied Probability, Chapman and Hall.



Taylor S.J. (1986), *Modelling Financial Time Series*, John Wiley & Sons.

Taylor, S.J., Xu X., (1997), The Incremental Volatility Information in One Million Foreign Exchange Quotations, *Journal of Empirical Finance*, 4, 317-340.

Thomakos D. Wang, T. D., (2003), Realized volatility in the futures markets, *Journal of Empirical Finance*, 10 (2003) 321-353

Wasserfallen, W., Zimmermann, H. (1985), The behavior of intraday exchange rates, *Journal of Banking and Finance*, 9, 55-72.

Wood, R., (2000), Market Microstructure Research Databases: History and Projections, *Journal of Business & Economic Statistics*, 18(2), 140-145.

Zakoian, J. M. (1990), *Threshold Heteroskedastic Models*, CREST, INSEE, Paris.

Zakoian, J. M. (1994), Threshold Heteroskedastic Models, *Journal of Economics Dynamics and Control*, 18, 931-944.



